

Aberystwyth University

Time and Cost Analysis for Repository Deposit

Payne, Hannah

Published in:
ALISS Quarterly

Publication date:
2011

Citation for published version (APA):

Payne, H. (2011). Time and Cost Analysis for Repository Deposit: The Welsh Repository Network Mediated Deposit Bureau. *ALISS Quarterly*, 6(3), 10-13.

Document License CC BY-NC-ND

General rights

Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400
email: is@aber.ac.uk

Time and cost analysis for repository deposit: the Welsh Repository Network Mediated Deposit Bureau

Hannah Payne, Repository Support Officer, Welsh Repository Network (WRN)

Introduction

As part of the Welsh Repository Network (WRN) ¹ Enhancement Project (WRN-EP)² an experimental service, the Mediated Deposit Bureau (MDB), was used to investigate the effect of outsourcing repository deposit on institutional repository services.

Working with a selection of pilot institutions from WRN partners, the project team based at Aberystwyth University took responsibility for the deposit of a number of identified collections. The objectives of the MDB were to:

1. Directly increase the number of deposited items within each repository.
2. Identify per unit costs, in both time and money, for repository item deposit.
3. Engage institution staff with the repository deposit process in an attempt to embed the processes involved within their working cycle and further increase deposit.
4. Encourage further deposit by academic staff within each repository through the presence of new repository deposits.

It is the second objective which provides the focus for this paper.

The full set of services the WRN team provided within the MDB was: record creation; full-item deposit permission checks and item deposit where possible; template requests for full-items and/or further permissions from authors or copyright holders. The WRN team did not approach authors or copyright holders to make deposit requests on an institution's behalf, however. It was felt more appropriate for an institution to do this as one, this would engage institution staff with the repository deposit process and two, would probably result in a more positive response.

Taking the idea forward

Three pilot institutions participated in the service: Swansea Metropolitan University (SMU); University of Wales Institute Cardiff (UWIC); University of Wales Newport (Newport). Each institution identified discrete collections of items they wished to submit for processing within the MDB. For the purposes of this paper however, the deposit of only the particular collections detailed below will be discussed.³

SMU Collection

SMU identified the research outputs submitted to their 2008 Research Assessment Exercise (RAE) return for deposit. The WRN team were provided with a PDF copy of the return which included full metadata records for each of the 131 items. The full set of services offered within the MDB was carried out. Records were prepared externally to the repository as the bulk-upload method for

¹ <http://www.wrn.aber.ac.uk/>

² <http://www.jisc.ac.uk/whatwedo/programmes/inf11/sue2/wrnep.aspx>

³ A full report on the MDB along with individual case studies will be available via the WRN website (<http://www.wrn.aber.ac.uk/>)

deposit to the institution's DSpace system⁴ was to be utilised due to the relatively high number of items.

UWIC Collection

The collection from UWIC consisted of the original computer files of a collated group of research outputs held in the School of Art and Design's SharePoint site. The WRN team were tasked only with the creation of repository records for the items held. In total, 10 of the stored items were identified as suitable for repository deposit and submitted directly via the DSpace deposit interface. Screen capture software was used to record each submission.⁵

Newport Collection

Staff at Newport had collated citation lists detailing outputs produced by academics from each of the University's academic schools. A record for each of the cited outputs was created and deposit check services carried out. It is those items identified for deposit from the School of Art, Media and Design list which are discussed in this paper. Again, deposit was made directly via the DSpace repository's submission interface, recorded by screen capture software.

Results

Institution	Total time taken (hours: minutes: seconds)	Cost in £ at £14.03 per hour	No. of items	Average time taken per item (minutes: seconds)	Cost in £ per item
SMU	33: 50: 00	474.68	130 (timed)	15: 37	3.65
UWIC	1: 45: 46	24.79	9	11: 45	2.75
Newport	00: 47: 32	11.11	4	11: 53	2.78

Table 1: Table to show the total amount of time taken with each MDB collection; the cost of that time based on the hourly rate of a Grade 6 Repository Support Officer; the average time taken per item; the average cost for processing each item in a given collection.

As can be seen from the data above, the average time taken per item was very similar between the UWIC SharePoint items and the Newport Art citation items (records for both sets were added directly via the repository interface). There was a difference of 8 seconds between the two average times with Newport records taking the longer length of time to complete. Deposit checks were performed on these items, where no such check was performed on the UWIC items. Also, it was possible to locate and upload a full-text copy of an article processed within the Newport collection which added time to the processing of this item and would have affected the average time calculated. However, it is difficult to imply that the deposit checks were the cause of the difference in time. The original metadata sources for the two collections were very different; data for the UWIC items was taken directly from the original files; data from the Newport items came from citations. The types of items processed within each set were slightly different also. The UWIC collection

⁴ Documentation for this process is available via the online DSpace System Documentation- Chapter 8: System Administration (http://www.dspace.org/1_6_0Documentation/ch08.html#N158A5)

⁵ A video of the on-screen processes within each submission were recorded via SMART Notebook Screen Capture software (<http://www.smarttech.com/>).

consisted of Articles, Presentations, a Book chapter and an image. The Newport collection consisted of three Presentations and one Article.

In contrast, the average time taken per item within the SMU collection is almost 4 minutes longer than the Newport or UWIC averages, at 15 minutes 37 seconds. This value included the length of time taken for bulk-upload preparation and deposit. This is a valid addition as it is important to consider the length of time taken to take the original metadata source to a submitted repository record. It can be said then that within the MDB it took a greater length of time to process and prepare items for bulk-upload than to process and deposit items directly via the repository.

Comparison with previous studies

Only one study could be identified that had considered the length of time necessary to make a repository deposit. Utilising repository log files, based on 260 repository deposits, Carr & Harnad (2005)⁶ calculated the average time taken for item metadata completion as 10 minutes 40 seconds. This is over a minute less than the shortest average time determined within the MDB for direct repository deposit. However, as highlighted by Joint (2006)⁷, items within this particular study were self-archived and the extent of the metadata entered may have been minimal. An objective for the WRN team was to deliver as full a metadata record as possible for each item deposited via the MDB.

A number of papers have suggested figures for the cost of repository deposit per item^{8, 9, 10, 11}. Only two however, have calculated this cost on the length of time taken for deposit^{12, 13}, the others calculating the cost by dividing the number of items present with the total running and set-up costs of the repository.

The Houghton report¹⁴ estimated the cost per deposit at £9.35, based on an average deposit time of 10 minutes at an average academic salary plus on-costs. This figure is over £5 more than the highest

⁶ Carr, L. and Harnad, S. (2005). *Keystroke Economy: A Study of the Time and Effort Involved in Self-Archiving*. Technical Report, ECS, University of Southampton. Available: <http://eprints.ecs.soton.ac.uk/10688/>

⁷ Joint, N. (2006). *Institutional repositories, self-archiving and the role of the library*. Library Review. 55(2), pp. 81-84.

⁸ Bevan, S. J. (2007). *Developing an institutional repository: Cranfield QJEprints- a case study*. OCLC Systems & Services: International digital library perspectives. 23(2), pp. 170-182.

⁹ Xia, J. & Sun, L. (2007). *Factors to assess self-archiving in institutional repositories*. Serials Review. 33(2), pp. 73-80. Available: <http://dx.doi.org/10.1016/j.serrev.2006.09.002>

¹⁰ Houghton, J., Rasmussen, B., Sheehan, P., Oppenheim, C., Morris, A., Creaser, C., Greenwood, H., Summers, M. & Gourlay, A. (2009). Economic implications of alternative scholarly publishing models: Exploring the costs and benefits. JISC Report. Available:

<http://www.jisc.ac.uk/media/documents/publications/rpteconomicoapublishing.pdf>

¹¹ Ayris, P., Davies, R., McLeod, R., Miao, R., Shenton, H. & Wheatley, P. (2008). The LIFE² final project report. Available: <http://discovery.ucl.ac.uk/11758/1/11758.pdf>

¹² Houghton, J., Rasmussen, B., Sheehan, P., Oppenheim, C., Morris, A., Creaser, C., Greenwood, H., Summers, M. & Gourlay, A. (2009). Economic implications of alternative scholarly publishing models: Exploring the costs and benefits. JISC Report. Available:

<http://www.jisc.ac.uk/media/documents/publications/rpteconomicoapublishing.pdf>

¹³ Ayris, P., Davies, R., McLeod, R., Miao, R., Shenton, H. & Wheatley, P. (2008). The LIFE² final project report. Available: <http://discovery.ucl.ac.uk/11758/1/11758.pdf>

¹⁴ Houghton, J., Rasmussen, B., Sheehan, P., Oppenheim, C., Morris, A., Creaser, C., Greenwood, H., Summers, M. & Gourlay, A. (2009). Economic implications of alternative scholarly publishing models: Exploring the costs

cost calculated within the MDB. This difference can be explained however, as the calculations within the MDB did not include on-costs, considering only the hourly rate of the Repository Support Officer. Also, it is assumed that the average salary value used within the Horton report would have been higher than the one used within the MDB.

As part of the LIFE² Project¹⁵ case studies from three repositories of SHERPA-LEAP¹⁶ (Goldsmiths, University of London; Royal Holloway, University of London; University College London (UCL)) detailed the cost of the full preservation lifecycle, as set out within the LIFE Model v1.1¹⁷, per repository deposit. The costs for the Acquisition, Ingest and Access stages of deposit were calculated between £14.70 at UCL up to £23 at Goldsmiths. These can be seen as significantly higher than the costs calculated within the MDB. Again however, the salary levels used within the calculations were different to the one used within the MDB and in some cases the costs within the case studies included the time of more than one member of staff. No estimate of times taken was provided however, making comparison difficult.

Conclusion

Although there have been a small number of studies considering the time or cost of repository deposit none can be seen to have considered both. Through the MDB, data for the average time taken for repository deposit through both the system interface and via the bulk-upload method has been derived. Staff costs for deposit based on WRN staff salary have also been calculated. However, it is possible for others to estimate the cost of their staff time using this data also.

and benefits. JISC Report. Available:

<http://www.jisc.ac.uk/media/documents/publications/rpteconomicoapublishing.pdf>

¹⁵ <http://www.life.ac.uk/2/>

¹⁶ <http://www.sherpa-leap.ac.uk/>

¹⁷ <http://discovery.ucl.ac.uk/4831/>